Application of Data Mining Techniques in Health Fraud Detection

Rekha Pal and Saurabh Pal
VBS Purvanchal University
Jaunpur, U.P., India
E-mail: drsaurabhpal@yahoo.co.in

ABSTRACT- Health smart card is like ATM card which provide cash benefits to patient through insurance company for hospital and medical benefits without expending money from the patient at the time of need. But now-a-days, fraud is done using the health smart card as few patients does not know the real cost of the treatment, so doctor take more payment and benefits through health smart card and generate fraud cash benefits. In this paper we proposed health fraud detection using different data mining techniques with the help of ID3, J48, Naïve Bayes. We have presented better accurate data by classifying all observations of fraud.

Keywords
Fraud detection; Classification: ID3, J48 and Naive Bayes; BPL.

1. INTRODUCTION

Fraud will always be a problem for many insurance companies. Data mining techniques can minimize some of these losses by making use of the collections of customer data, particularly in insurance, credit card, and telecommunications companies

Health care fraud, based on the definition of the NHCAA (National Health Care Anti-fraud Association), is an intentional deception or misrepresentation made by a person or an entity, with the knowledge that the deception could result in some kinds of unauthorized benefits to that person or entity.

In recent years, systems for processing electronic claims have been increasingly implemented to automatically perform audits and reviews of claims data. These systems are designed for identifying areas requiring special attention such as erroneous or incomplete data input, duplicate claims, and medically non-covered services. Although these systems may be used to detect certain types of fraud, their fraud detection capabilities are usually limited since the detection mainly relies on pre-defined simple rules specified by domain experts.

In order to assure the batter operation of a health care insurance system, fraud detection mechanisms are imperative, but highly specialized domain knowledge is required. Furthermore, well-designed detection policies, able to adapt to new trends acting simultaneously as prevention measures, have to be considered. Data mining which is part of an iterative process called knowledge discovery in databases (KDD) [9] [10] can assist to extract this knowledge automatically. It has allowed better direction and use of health care fraud detection and investigative resources by recognizing and quantifying the underlying attributes of fraudulent claims, fraudulent providers, and fraudulent beneficiaries [11]. Automatic fraud detection helps to reduce the manual parts of a fraud screening/checking process becoming one of the most established industry/government data mining applications [7].

A health smart card is a type of plastic card that contains an embedded computer chip–either a memory or microprocessor type–that stores and transacts data. This data is usually associated with either value or information, or both and is stored and processed within the card's chip. The card data is transacted via a card reader that is part of a computing system. Systems that are enhanced with smart cards are in use today throughout several key applications, including healthcare, banking, entertainment, and transportation. All patients do not know about all facility of his card and cost of treatment, so some doctor easily take financial benefit of the health smart card easily and make fraud with patients. Every "below poverty line" (BPL) family holding a yellow ration card pays 30 registration fee to get a biometric-enabled smart card containing their fingerprints and photographs. This enables them to receive inpatient medical care of up to 30,000 per family per year in any of the empanelled hospitals.

2. DATA MINING

Data mining is a term from computer science. Sometimes it is also called knowledge discovery in databases (KDD). Data mining is about finding new information in a lot of data. The information obtained from data mining is hopefully both new and useful. In many cases, data is stored so it can be used later. The data is saved with a goal. Saving information makes a lot of data. The data is usually saved in a database. Finding new information that can also be useful from data is called data mining.

DIFFERENT TYPES OF DATA MINING TECHNIQUES

There are lots of different types of data mining techniques for getting new information.

DECISION TREES

A decision tree is a classifier expressed as a recursive partition of the instance space. The decision tree consists of nodes that form a rooted tree, meaning it is a directed tree with a node called "root" that has no incoming edge. All other nodes have exactly one incoming edge. A node with outgoing edges is called an internal or test node. All other nodes are called leaves (also known as terminal or decision nodes). In a decision tree, each internal node splits the instance space into two or more sub-spaces according to a certain discrete function of the input attributes value. In the simplest and most frequent case, each test considers a single attribute, such that the instance space is partitioned according to the attribute's value. In the case of numeric attributes, the condition refers to a range .Each leaf is assigned to one class representing the most appropriate target value. Alternatively, the leaf may hold a probability vector indicating the probability of the target attribute having a certain value. Instances are classified by navigating them from the root of the tree down to a leaf, according to the outcome of the tests along the path describes a decision tree that reasons. Decision tree algorithms ID3, J48 and NB Tree can be applied on large amount of data and valuable predictions can be produced. These predictions evaluate future behavior of problem. Decision tree are preferred because they can evaluate information more accurately than other methods. These three algorithms shows the probability of input attributes respectively.

ID3 decision tree algorithm introduced in 1986 by Quinlan Ross [1]. It is based on Hunts algorithm. The tree is constructed in two phases. The two phases are tree building and pruning. ID3 uses information gain measure to choose the splitting attribute. It only accepts categorical attributes in building a tree model. It does not give accurate result when there is noise. To remove the noise pre-processing technique has to be used. To build decision tree, information gain is calculated for each and every attribute and select the attribute with the highest information gain to designate as a root node. Label the attribute as a root node and the possible values of the attribute are represented as arcs. Then all possible outcome instances are tested to check whether they are falling under the same class or not. If all the instances are falling under the same class, the node is represented with single class name, otherwise choose the splitting attribute to classify the instances. Continuous attributes can be handled using the ID3 algorithm by discretizing or directly, by considering the values to find the best split point by taking a threshold on the attribute values. ID3 does not support pruning.

Decision trees and decision rules introduced by Apte and Weiss [2] On attribute values in decision tree generation introduced by Fayyad [3].

J48 classification by decision tree induction decision tree-leaf nodes represent class labels or class distribution

1. A flow-chart-like tree structure internal node denotes a test.
2. On an attribute branch represents an outcome of the test.
3. Decision tree generation consists of two phase's tree.
4. Construction at start, selected attributes tree pruning.
5. Identify and remove branches that reflect noise or outliers.

The Naïve Bayes [4] Classifier technique is particularly suited when the dimensionality of the inputs is high. Despisimplicity, Naive Bayes can often outperform more sophisticated classification methods. Naïve Bayes model identifies the characteristics of dropout students. It shows the probability of each input attribute for the predictable state. A Naive Bayesian classifier is a simple probabilistic classifier based on applying Bayesian theorem (from Bayesian statistics) with strong (naive) independence assumptions. By the use of Bayesian theorem we can write.

$$P(c_i \mid x) = \frac{P(x \mid c_i)P(c_i)}{P(x)}$$

We preferred Naive Bayes implementation because: Simple and trained on whole (weighted) training data

- Over-fitting (small subsets of training data) protection.
- Claim that boosting "never over-fits" could not be maintained.
- Complex resulting classifier can be determined reliably from limited amount of Data.

Naive Bayesian tree consists of [5] classification and decision tree learning. An NB Tree classification sorts to a leaf and then assigns a class label by applying a Naïve Bayes on that leaf. The steps of NB Tree algorithm are:

(a) At each leaf node of a tree, a Naive Bayes is applied.

(b) By using Naive Bayes for each leaf node, the instances are classified.

(c) As the tree grows, for each leaf a Naive Bayes is constructed

The Weka workbench contains a collection of visualization tools and algorithms for data analysis and predictive modeling, together with graphical user interfaces for easy access to this functionality. It is portable and platform independent because it is fully implemented in the Java programming language and thus runs on almost any modern computing platform. Weka has several standard data mining tasks, data preprocessing, clustering, classification, association. The easiest way to do this is simply to download the template, and replace the content with your own material.

3. RELATED WORK-

Al Lin and Yes [6] discussed that the following practices is helpful to adjust the accuracy and coverage ratio of the model, which include the removal of unnecessary data, making data more representative and the preprocessing of missing data.

(a) Removal of unnecessary data: The pattern of question inquiry or thinking can be changed so as to remove needless samples. For example, compared with all applicants or people with or without payment, the relative minority can be removed preferentially when the model is constructed.

(b) Data representative: Where appropriate, advanced data processing methods can be added (such as existing data generated from some statistics) to reinforce the rationality of data interpretation and summarization and to make the data abundant.

(c) Processing of data missing: The missing of some data in the database will bring about difficulties in analyzing. Although some methods can be adopted to supplement the data, the data are not all real and may be inaccurate at times. Therefore, sometimes rules for removing or methods and principles for supplement the missing data can be thought about.

Ashathna and Madan [7] discussed that educational data mining is an area full of opportunities where one can know how to improve student success and institute effectiveness and make it efficient. It is a systematic process which offers firms ability to discover hidden patterns so that all the problems coming forward can be solved and customized and proper decision making can be done. It is concerned with determining the pattern from the hidden information so that the different student's behavior can be identified and according to those necessary steps can be taken and we can know the faculty feedback so as to perform any changes in the work structure or working schedule.

Derrig[8] discussed that measurement, detection, and deterrence of fraud are advanced through statistical models, intelligent technologies are applied to informative databases to provide for efficient claim sorts, and strategic analysis is applied to property-liability and health insurance situations.

Phua, Lee, Smith and Gayler [9] discussed that fraud and corruption in health care industry can be grouped in illicit activities associated to affiliates, medical professionals, staff and manager, and suppliers. Although fraud may not necessarily lead to direct legal consequences, it can become a critical problem for the business if it is very prevalent and if the prevention procedures are not failsafe.

Sokol, Garcia, Rodriguez, West and Ohnson [10] discussed that knowledge discovery in databases (KDD) can assist to extract this knowledge automatically. It has allowed better direction and use of health care fraud detection and investigative resources by recognizing and quantifying the underlying attributes of fraudulent claims, fraudulent providers, and fraudulent beneficiaries.

Pflaum and Rivers [11] discussed that in several countries fraudulent and abusive behavior in health insurance is a major problem. Fraud in medical insurance covers a wide range of activities in terms of cost and sophistication.

Yang , Hwang and Opit [12][13] discussed that health insurance systems are either sponsored by governments or managed by the private sector, to share the health care costs in those countries.

Major and Riedinger [14] discussed that a set of behavioral rules based on heuristics and machine learning are used for performing and scanning a large population of health insurance claims in search of likely fraud.

keuchi, Williams, and Milne [15] discussed that in an on-line discounting learning algorithm to indicate whether a case has a high possibility of being a statistical outlier in data mining applications such as fraud detection is used for identifying meaningful rare cases in health insurance pathology data from Australia's Health Insurance Commission (HIC).

Tennyson and Salsas [16] discussed that in automobile insurance fraud prevention studies case, it results distinguish between two kinds of results, suspected fraud cases and reasonable. Prediction method use Logistic Regression.

In Bureau of National Health insurance [17] discussed that data mining is a process, one of the procedures of knowledge discovery, to discover useful models in the data probably to be used in the future, which have never been seen previously and are easily to be understood.

Weisberg and Derrig [18] discussed that to use these characteristics in the development of insurance fraud reasoning system. These characteristics will be reviewed with an experienced expert assess of the rules of the medical insurance fraud prevention knowledge, for comparison. The medical insurance fraud characteristics include: Damage level insufficient information, suspected diagnosis of proof, insured low willingness to cooperate and Cause of the accident unreasonable.

Becker, Kessler and McClellan [19] discussed that identify the effects of fraud control expenditures and hospital and patient characteristics on up coding, treatment intensity and health outcomes in the Medicare and Medicaid programs.

Cox [20] discussed that applied a fraud detection system based on fuzzy logic for analyzing health care provider claims.

Yadav and Pal [21] conducted a study using classification tree to predict student academic performance using students' gender, admission type, previous schools marks, medium of teaching, location of living, accommodation type, father's qualification, mother's qualification, father's occupation, mother's occupation, family annual income and so on. In their study, they achieved around 62.22%, 62.22% and 67.77% overall prediction accuracy using ID3, CART and C4.5 decision tree algorithms respectively.

In another study Yadav et al. [22] used students' attendance, class test grade, seminar and assignment marks, lab works to predict students' performance at the end of the semester with the help of three decision tree algorithms ID3, CART and C4.5. In their study they achieved 52.08%, 56.25% and 45.83% classification accuracy respectively.

Pal and Pal [23] conducted study on the student performance based by selecting 200 students from BCA course. By means of ID3, c4.5 and Bagging they find that SSG, HSG, Focc, Fqual and FAIn were highly correlated with the student academic performance.

Vikas Chaurasia et al. [24] used CART (Classification and Regression Tree), ID3 (Iterative Dichotomized 3) and decision table (DT) to predict the survivability for Heart Diseases patients.

Bhardwaj and Pal [25] conducted study on the student performance based by selecting 300 students from 5 different degree college conducting BCA (Bachelor of Computer Application) course of Dr. R. M. L. Awadh University, Faizabad, India. By means of Bayesian classification method on 17 attributes, it was found that the factors like students' grade in senior secondary exam, living location, medium of teaching, mother's qualification, students other habit, family annual income and student's family status were highly correlated with the student academic performance.

In this paper we proposed health smart card fraud detection different data mining techniques with weka tool we have presented better accurate data by classifying all observations of fraud.

4. METHODOLOGY

4.1 Data Preparation

A medical post-operative claim charges through smart card payment involves the participation of patient admitted in a Heart care hospital. Claim accepted and investigated by ICICI Lombard and oriental insurance company. A BPL card holder patient admitted in Heart care hospital for heart treatment purpose her Guardian made payment charge by doctor through BPL card the charged claim for payment by hospital authority is 10,000. But the actual amount is only 7,000. The fraud conducted by beneficiary hospital authority is approximately 3,000. Now for performing classification of BPL using several standard data mining tasks, data preprocessing, clustering, classification, association and tasks are needed to be done. The database is designed in MS-Excel and MS word 2010 database management system to store the collected data. The data is formed according to the required format and structures. Further, the data is. converted to ARFF (Attribute Relation File Format) format to process in WEKA. An ARFF file is an ASCII text file that describes a list of instances sharing a set of attributes. ARFF files were developed by the Machine Learning Project at the department of computer science of the university of Waikato for use with the Weka machine learning software which give the solutions by algorithms tools.

4.2 Data selection and transformation

The variables used in the computational technique to identify the fraud claim or none fraud auto motive in insurance claims.The Kappa statistic (or value) is a metric that compares an observed accuracy with an expected accuracy (random chance).The kappa statistic is used not only to evaluate a single classifier, but also to evaluate classifiers amongst themselves computation of observed accuracy and expected accuracy is integral to comprehension of the kappa statistic, and is most use of easily illustrated through a confusion matrix.

TABLE 1.The variables used in the computational technique.

PROPERTY	**DESCRIPTION**	
Source	Heart care hospital	
Claim type	Post-operative claim charges through smart card payment	
Sample size	61 Total:2 fraud and 59 LGT examine by ICICI Lombard and oriental insurance company.	
Dependable variable		
Fraud(0)	Claim not accepted	
Lgt(1)	Claim accepted	
Explanatory variable	Value	Description
Is	{0-fraud}	Not accepted claim of smart card by investigated surveyor of ICICI Lombard insurance company.
	{1-legle}	Accepted claim of smart card by investigator surveyor of ICICI Lombard insurance company.
Date gap	{0-zeroday,1-threedays,2-sixdays,3-ninedays,4-twalvedays}	Time difference (in day) between insurance claim and the policy report being filled.
Bnp	{0-new policy not found,1-new policy found}	BNP indicates Boolean value for new policy yes (1) or (0).
Clmamt	{0-five thousands,1-ten thousands,2-fifteen thousands,3-twenty thousands,4-twenty five thousands}	Claim amount as a static value.
Fyrs	{0-Year,1-Oneyear,2->Oneyear,3-=Twoyears,4->Two years}	Financial year in which complained.
Clms	{0-Zero,1-One,2-Two,4-Four}	Total number of claims the claimant has filled with in insurance company.

The variables used in the computational technique to identify the fraud claim or none fraud auto motive in insurance claims.The Kappa statistic (or value) is a metric that compares an observed accuracy with an expected accuracy (random chance).The kappa statistic is used not only to evaluate a single classifier, but also to evaluate classifiers amongst themselves computation of observed accuracy and expected accuracy is integral to comprehension of the kappa statistic, and is most use of easily illustrated through a confusion matrix.

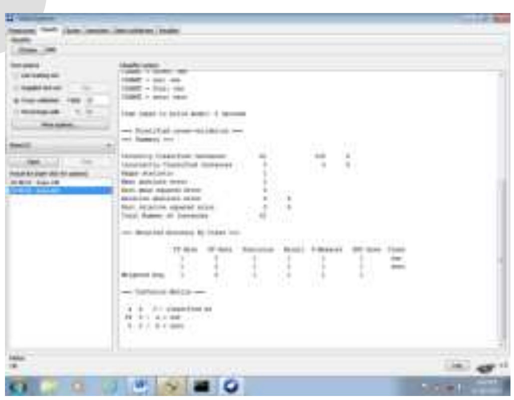

Figure.1. Instances classified by J48.

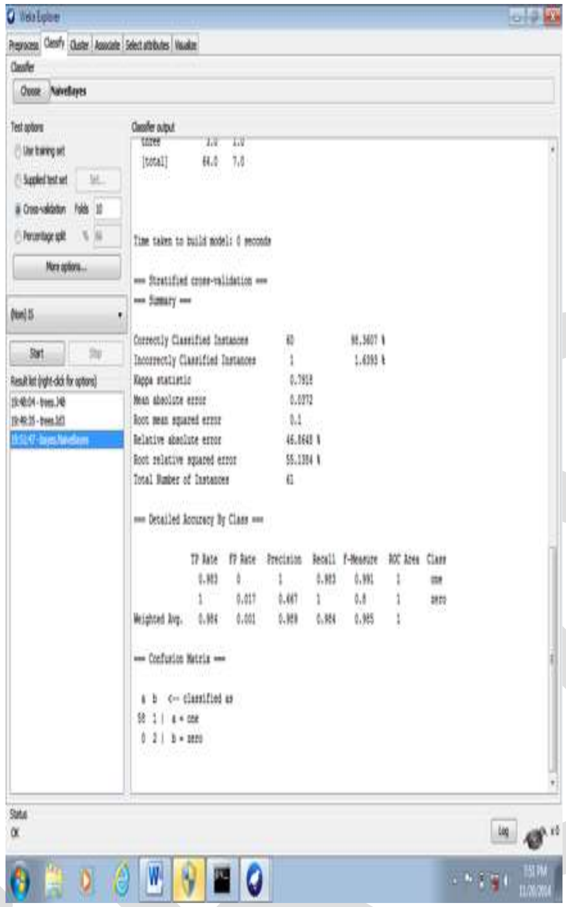

Figure-2. Instances classified by ID3.

Figure -3. Instances classified by NAÏVE BAYES

General, Positive = identified and negative = rejected. Therefore:

- True positive = correctly identified
- False positive = incorrectly identified
- True negative = correctly rejected
- False negative = incorrectly rejected

A confusion matrix, also known as a contingency table or an error matrix is a specific table layout that allows visualization of the performance of an algorithm, typically a supervised learning one (in unsupervised learning it is usually called a matching matrix). Each column of the matrix represents the instances in a predicted class, while each row represents the instances in an actual class.

4.3 Implementation of Data Mining

The paper presents an approach to classifying health smart card in order to predict their amount based on features extracted from logged data in a hospital. They design, implement, and evaluate a series of pattern classifiers and compare performance of an online patient's dataset. Classifiers were used to declare surety of health card holder. They used the J48, ID3 and Naïve Bayes algorithms to improve the prediction accuracy. This method is of considerable usefulness in identifying patient in very large real data set and allow investigator to provide appropriate investigate in a timely manner.

4.4 Results and Discussion

The proposed techniques which are include in WEKA tool, Decision Tree and Bayes techniques. Data mining tools are software components. The proposed tool that will be applied is WEKA software tool because it is support several standard data mining tasks. WEKA is a collection of machine learning algorithms for solving real-world data mining problems. It is written in Java and runs on almost any platform. The proposed technique that will be applied in this paper is Decision Tree (J48, ID3 and NAÏVE BAYES) because it is powerful classification algorithms. From the table-2. it is clear that ID3 Kappa static value observed give the greater value compare to J48 and NB algorithms and ID3 give the less error compare to J48 and NB algorithms.ID3 take 0.2sec in process completion but J48 and NB take 0 sec.

Table -2 Evaluation on test split

		J48	ID3	NAÏVE BAYES
TP rate	LGT	1	1	0.98
	FRAUD	0	1	1
	WEIGHTED	0.96	1	0.98
FP rate	LGT	1	0	0
	FRAUD	0	0	0.02
	WEIGHTED	0.96	0	0
PRECISION	LGT	0.96	1	1
	FRAUD	0	1	0.67
	WEIGHTED	0.94	1	0.99
RECALL	LGT	1	1	0.98
	FRAUD	0	1	1
	WEIGHTED	0.97	1	0.98
F-MEASURE	LGT	0.98	1	0.99
	FRAUD	0	1	0.8
	WEIGHTED	0.95	1	0.99
ROC	LGT	0.09	1	1
	FRAUD	0.09	1	1
	WEIGHTED	0.09	1	1

Table-3. Detailed Accuracy by Class

Detailed	J48	ID3	Naïve Bayes
Correctly classified Instances	59	61	60
Incorrectly classified instances	2	0	1
Kappa static	0	1	0.79
Mean absolute error	0.06	0	0.04
Root mean squared error	0.18	0	0.1
Relative absolute error	80.71 %	0%	46.86 %
Root relative square error	99.67 %	0%	55.13 %
Time taken (Second)	0	0.02	0
Total number of instances	61	61	61

From table 3 it is clear that ID3 give more correctly classified compare to J48 and NB algorithms and ID3 provide greater value of weighted avg. compare to J48 and NB algorithms.

Table-4. Confusion Matrix

Confusion Matrix(J48)	Confusion Matrix(ID3)	Confusion Matrix(Naïve Bayes)
59 0 | a = one	59 0 | a = one	58 1 | a = one
2 0 | b = zero	0 2| b = zero	0 2 | b = zero

Each column of the matrix represents the instances in a predicted class, while each row represents the instances in an actual class. In table4.confusion matrix J48 shows 59 correctly classified and 2 none correctly classified another side 0 none correctly classified and 0 correctly classified arise. Confusion matrix In ID3 represents 59 correctly classified and 0 none correctly classified another side 0 none correctly classified and 2 correctly classified. In NB 58 correctly classified 0 none correctly classified another side 1 none correctly classified and 2 correctly classified. It is clear from analysis correctly classified in ID3 is total instances 61 is greater value compare to other J48 and NB algorithms.

5. CONCLUSION

In WEKA, all data is considered as instances attributes in the data. For easier analysis and evaluation, simulation results are partitioned into several sub items. In the first part, correctly and incorrectly classified instances will be partitioned in numeric and percentage value, and subsequently, Kappa statistics, mean absolute error and root mean squared error will be at a numeric value only. ID3 time taken to build model: 0.02 seconds and test mode: 10-fold cross-validation .Here weka computes all required parameters on given instances with the classifiers' respective accuracy and prediction rate. Based on table2 we can clearly see that the highest accuracy of ID3 is 100% and the lowest accuracy of J48 is 96.7213% so decision tree ID3 is the best in the three respective algorithms as it is more accurate.

REFERENCES:

[1] J. R. Quinlan, C4.5: Programs for machine learning, Morgan Kaufmann, San Francisco, 1993. .

[2] C. Apte and S. Weiss. Data mining with decision trees And decision rules. Future Generation Computer Systems, 13,1997.

[3] U. M. Fayyad. Branching on attribute values in decision tree generation. In Proc. 1994 AAAI Conf.Pages601-606, AAAI Press, 1994.

[4] Harry Zhang "The Optimality of Naive Bayes". FLAIRS2004 conference.

[5] Yumin Zhao, Zhendong Niu_ and Xueping Peng, "Research on Data Mining Technologies for Complicated Attributes Relationship in Digital Library Collections" "Applied Mathematics and Information Sciences, An International Journal", Appl. Math. Inf. Sci. 8, No. 3, 1173-1178 (2014).

[6] Kuo-Chung Lin, Ching-Long Yeh "Use of Data Mining Techniques to Detect Medical Fraud in Health Insurance" International Journal Of Engineering and Technology Innovation, vo no. , 201, pp. 126-137

accepted 26 February 2012.

[7] Anansha Asthana, Komal R. Madan, "Educational Data Mining" in National Student Conference on Advances in Electrical and Informatio Communication Technology"AEICT-2014.

[8] Derrig R. A., "Insurance fraud," Journal of Risk and Insurance, vol. 69.3, pp. 271-287, 2002.

[9] Phua, Lee, Smith, and Gayler, "A comprehensive survey of data mining-based fraud detection research", Artificial Intelligenc Review, submitted, 2005.

[10] Sokol, Garcia, Rodriguez, West and Johnson, "Using data mining to find fraud in HCFA health care claims,"Top Health Information Management, vol. 22, no. 1, pp. 1-13, 2001.

[11] Pflaum and J. S. Rivers, "Employer strategies to combat health care plan fraud," Benefits quarterly, vol. 7, no. 1, pp. 6-14, 1991.

[12] W. S. Yang and S. Y. Hwang, "A process-mining framework for the detection of healthcare fraud and abuse," Expert Systems with Application Article in Press, corrected proof, 2005.

[13] L. J. Opit, "The cost of health care and health insurance in Australia: Some problems associated with the fee for-service," Soc. Sci. Med.vol. 18, no. 11, 967–972, 1984.

[14] J. Major, and D. Riedinger, "EFD: A hybrid knowledge/statistical based system for the detection of fraud," Journal of Risk and Insurance 69(3), pp. 309-324, 2002.

[15] K. Yamanishi, J. Takeuchi, G. Williams, and P. Milne, "On-line Unsupervised outlier detection using finite mixtures with discounting learning algorithms," In Data Mining and Knowledge Discovery vol. 8, pp. 275–300, 2004.

[16] Sharon Tennyson and Pau Salsas, "Patterns of Auditing in Markets with Fraud: Some Empirical Results from Automobile Insurance," Working Paper, 2000.

[17] Bureau of National Health insurance, Codes and scopes of classifications of diseases, National association of insurance commissioners,2014.

[18] Weisberg, H.I. and Derrig, R.A., "AIB Claim Screening Experiment Final Report," Insurance Bureau of Massachusetts, 1998.

[19] D. Becker, D. Kessler and M. McClellan, "Detect in Medicare abuse," Journal of Health Economics, vol. 24, pp. 189–210, 2002.

[20] E. Cox, "A fuzzy system for detecting anomalous behaviors in healthcare provider claims, Goonatilake, S. and Treleaven, P (eds.Intelligent Systems for Financial and Business, pp. 111-134.John Wiley and Sons Ltd., 1995.

[21] S. K. Yadav, Pal., S. 2012. Data Mining: A Prediction for Performance Improvement of Engineering Students using Classification, World of Computer Science and Information Technology (WCSIT), 2(2), 51-56.

[22] S. K. Yadav, B. K. Bharadwaj & Pal, S. 2011. Data Mining Applications: A comparative study for predicting students' performance, International journal of Innovative Technology and Creative Engineering (IJITCE), 1(12).

[23] A. K. Pal, and S. Pal, "Analysis and Mining of Educational Data for Predicting the Performance of Students", (IJECCE) International Journal of Electronics Communication and Computer Engineering, Vol. 4, Issue 5, pp. 1560-1565, ISSN:2278-4209, 2013.

[24] V. Chauraisa and S. Pal, "Early Prediction of Heart Diseases Using Data Mining Techniques", Carib.j.SciTech,,Vol.1, pp. 208-217, 2013.

[25] B.K. Bharadwaj and S. Pal. "Data Mining: A prediction for performance improvement using classification", International Journal of Computer Science and Information Security (IJCSIS), Vol. 9, No. 4, pp. 136-140, 2011.